

QUASI-IDENTITIES AND THE CAYLEY-HAMILTON QUASI-POLYNOMIAL

MATEJ BREŠAR AND ŠPELA ŠPENKO

ABSTRACT. A quasi-identity of the matrix algebra M_n is defined as a quasi-polynomial $P = P(X_1, \dots, X_m)$ such that $P(A_1, \dots, A_m) = 0$ for all $A_1, \dots, A_m \in M_n$. We consider the question whether every quasi-identity is a consequence of the quasi-identity arising from the Cayley-Hamilton theorem. Besides verifying that this holds in some special cases, we show that for an arbitrary quasi-identity P there exists a central polynomial c such that cP has this property. Some other properties of quasi-polynomials are also discussed, and the problem of describing the nonstandard solutions of more general functional identities is touched upon.

1. INTRODUCTION

A functional identity is an identical relation in a ring that, besides arbitrary elements (that appear in a similar fashion as in a polynomial identity), also involves arbitrary functions which are considered as unknowns; the goal is to describe their form (see [3]). One usually first finds the “obvious” solutions, i.e., those functions that satisfy a given functional identity for formal reasons, independent of the structure of the ring in question. These are called the *standard solutions*. A typical result says that either the standard solutions are also the only possible solutions or the ring has some special properties, like satisfying a polynomial identity of a certain degree related to the number of variables. To the best of our knowledge, *nonstandard solutions* have not been considered much so far. Therefore, the theory of functional identities has primarily served as a complement to the theory of polynomial identities, rather than its generalization. The purpose of this paper is to initiate the study of nonstandard solutions of functional identities, and in this way try to find tighter connections with polynomial identities.

It seems out of reach to consider nonstandard solutions of general functional identities, at least at this stage. Except for the final section, we will confine ourselves to quasi-identities, i.e., quasi-polynomials that vanish at all evaluations. Precise definitions of these notions and their basic properties will be given in the next two sections; speaking very roughly, however, one can think of a quasi-polynomial as of a noncommutative polynomial in which coefficients are not necessarily scalars but scalar-valued functions. Quasi-polynomials were introduced by Beidar and Chebotar in 2000 [1], and have since played a fundamental role in the theory of functional identities (in some papers, e.g. in [6], the authors call them Beidar polynomials). Standard solutions of quasi-identities can be very easily described: all coefficient functions must be 0 (cf. [3, Lemma 4.4]). The Cayley-Hamilton theorem gives rise to a basic example of a quasi-identity on the matrix algebra M_n with nonstandard solutions. The main theme of this

Key words and phrases. Functional identity, quasi-polynomial, quasi-identity, Cayley-Hamilton quasi-polynomial, T-ideal, trace identity, polynomial identity.

2010 *Math. Subj. Class.* Primary 16R60. Secondary 16R10, 16R30.

Supported by ARRS Grant P1-0288.

paper is the question whether every quasi-identity P of M_n is a consequence of this particular quasi-identity. One of the motivations for this problem is the well-known Helling-Procesi-Razmyslov theorem stating that the answer to such a question is positive for (somewhat similar) trace identities. We are unable to give a complete answer to our question, but do show that there exists a central polynomial $c \neq 0$ such that cP has the desired property (Theorem 5.3). The proof is based on the Helling-Procesi-Razmyslov theorem and on an auxiliary result (Theorem 4.2) that gives a useful supplement to the recent study of locally linearly dependent polynomials [4]. Furthermore, we obtain a definitive conclusion in the case where P is a multilinear quasi-identity of degree n (Theorem 6.1) or a quasi-identity in one indeterminate (Theorem 7.2). At the end of the paper we consider the nonstandard solutions of another special type of functional identities (Theorem 8.1).

Various problems on functional identities studied in [3] can be solved for quite general classes of rings. The study of nonstandard solutions, however, is of a different nature. We will deal exclusively with the algebra $M_n = M_n(F)$ of all $n \times n$ matrices over a field F . We also assume, without further mention, that $\text{char}(F) = 0$. It will be apparent from the proofs that some of results may hold for more general prime PI-algebras, and that the assumption on the characteristic is not always necessary. But we shall not elaborate on this here.

2. QUASI-POLYNOMIALS AND QUASI-IDENTITIES

We will define a quasi-polynomial in a slightly different way as in [1] and [3]. Our definition is not restricted to the multilinear situation, and, on the other hand, is adjusted for applications to the matrix algebra M_n .

We fix an integer $n \geq 2$, and set

$$\mathcal{C} := F[x_{ij}^{(k)} \mid 1 \leq i, j \leq n, k = 1, 2, \dots]$$

and

$$\mathcal{X} := \{X_k \mid k = 1, 2, \dots\}.$$

A *quasi-polynomial* is an element of the algebra

$$\mathfrak{Q}_n := \mathcal{C}\langle\mathcal{X}\rangle.$$

Thus, a quasi-polynomial is a polynomial in the noncommuting indeterminates X_k whose coefficients are ordinary polynomials in the commuting indeterminates $x_{ij}^{(k)}$. A quasi-polynomial P can be therefore written as

$$P = \sum \lambda_M M,$$

where M is a noncommutative monomial in the X_k 's and λ_M is a commutative polynomial in the $x_{ij}^{(k)}$'s. Of course, P depends on finitely many X_k 's and finitely many $x_{ij}^{(k)}$'s. We can therefore write

$$P = P(x_{11}^{(1)}, \dots, x_{nn}^{(1)}, \dots, x_{11}^{(m)}, \dots, x_{nn}^{(m)}, X_1, \dots, X_m)$$

for some m . We define the *evaluation* of P at an m -tuple $A_1, \dots, A_m \in M_n$, $P(A_1, \dots, A_m)$, by substituting A_k for X_k and $a_{ij}^{(k)}$ for $x_{ij}^{(k)}$, where $A_k = (a_{ij}^{(k)})$. If $P(A_1, \dots, A_m) = 0$ for all $A_1, \dots, A_m \in M_n$, then we say that P is a *quasi-identity* of M_n . It is convenient to use a more suggestive notation and write $\lambda_M(x_{11}^{(1)}, \dots, x_{nn}^{(1)}, \dots, x_{11}^{(m)}, \dots, x_{nn}^{(m)})$ as $\lambda_M(X_1, \dots, X_m)$, i.e., consider the polynomial λ_M as a polynomial function $M_n^m \rightarrow F$. Hence we can write P as

$P(X_1, \dots, X_m)$. Now we define a T -ideal of \mathfrak{Q}_n as an ideal \mathcal{I} such that if $P(X_1, \dots, X_m) \in \mathcal{I}$, then $P(H_1, \dots, H_m) \in \mathcal{I}$ for all $H_1, \dots, H_m \in \mathfrak{Q}_n$. The set \mathfrak{I}_n of all quasi-identities of M_n obviously forms a T -ideal of \mathfrak{Q}_n .

An obvious example of a quasi-identity of M_n is the quasi-polynomial arising from the Cayley-Hamilton theorem,

$$q_n = q_n(X_1) = X_1^n + \tau_1(X_1)X_1^{n-1} + \dots + \tau_n(X_1),$$

where $\tau_1(X_1) = -\text{tr}(X_1) = -(x_{11}^{(1)} + \dots + x_{nn}^{(1)})$, \dots , $\tau_n(X_1) = (-1)^n \det(X_1)$. As it is well-known, each $\tau_i(X_1)$ can be expressed as a \mathbb{Q} -linear combination of the products of $\text{tr}(X_1^j)$. Let $Q_n(X_1, \dots, X_n)$ denote a multilinear version of $q_n(X_1)$. Thus, for example,

$$Q_2(X_1, X_2) = X_1X_2 + X_2X_1 - \text{tr}(X_1)X_2 - \text{tr}(X_2)X_1 + \text{tr}(X_1)\text{tr}(X_2) - \text{tr}(X_1X_2).$$

Note that $Q_n(X_1, \dots, X_n)$ is symmetric, i.e., $Q_n(X_1, \dots, X_n) = Q_n(X_{\sigma(1)}, \dots, X_{\sigma(n)})$ for every permutation σ , and that $q_n(X_1) = \frac{1}{n!}Q_n(X_1, \dots, X_1)$. We will call Q_n the *Cayley-Hamilton quasi-polynomial*. This paper is primarily devoted to the following problem.

Problem 2.1. *Is every quasi-identity of M_n contained in the T -ideal of \mathfrak{Q}_n generated by Q_n ?*

In other words, we are asking whether the T -ideal \mathfrak{I}_n is generated by Q_n (here we may replace Q_n by q_n , as q_n and Q_n generate the same T -ideal).

The Cayley-Hamilton quasi-polynomial Q_n is also a basic example of a *trace identity*. We refer the reader to [9, Chapter 12] for an account on such identities. Informally, a trace polynomial can also be viewed as a quasi-polynomial $\sum \lambda_M M$, but such that every λ_M can be expressed as a linear combination of the products of $\text{tr}(X_{i_1} \cdots X_{i_m})$. However, there is an important difference between the notions of trace identities and quasi-identities. For instance, $\text{tr}(Q_n(X_1, \dots, X_n)X_{n+1})$ is a nontrivial trace identity, but a trivial quasi-identity. A well-known theorem, obtained independently by Helling [7], Procesi [10] and Razmyslov [11], states that the T -ideal of trace identities of M_n is generated by $Q_n(X_1, \dots, X_n)$ and $\text{tr}(Q_n(X_1, \dots, X_n)X_{n+1})$. In particular, since nonzero polynomial identities are nontrivial trace identities as well as nontrivial quasi-identities, this implies

Theorem 2.2. *Every polynomial identity of M_n is contained in the T -ideal of \mathfrak{Q}_n generated by Q_n .*

Thus, Problem 2.1 asks whether Theorem 2.2 still holds if one replaces “polynomial identity” by “quasi-identity”. Another problem, which is perhaps also of interest and we were unable to resolve, is to find a new proof of Theorem 2.2, independent of trace identities. This could help in a better understanding of polynomial identities and their relation with functional identities.

3. GENERIC INTERPRETATION

The algebra of quasi-polynomials \mathfrak{Q}_n contains the algebra of noncommutative polynomials $F\langle \mathcal{X} \rangle$. How much larger it is? Of course, this vague question may not have an accurate answer. Still, remarks in this section suggest that “substantially larger” may be a simplified answer.

Similarly, the ideal \mathfrak{I}_n of \mathfrak{Q}_n of quasi-identities of M_n contains the ideal $\text{id}(M_n)$ of $F\langle \mathcal{X} \rangle$ of polynomial identities of M_n . When studying $\text{id}(M_n)$, one often passes to the algebra of generic matrices, which is isomorphic to the relatively free algebra $F\langle \mathcal{X} \rangle / \text{id}(M_n)$. Therefore

it seems natural to ask what can we say about the factor algebra $\mathfrak{Q}_n/\mathfrak{I}_n$. Our first remark is that, unlike $F\langle\mathcal{X}\rangle/\text{id}(M_n)$, it is not a domain. This can be deduced from Lemma 3.3 below, but let us, nevertheless, give a simple concrete example.

Example 3.1. Note that none of

$$P_1 = x_{12}^{(2)} X_1 - x_{12}^{(1)} X_2 + x_{12}^{(1)} x_{22}^{(2)} - x_{22}^{(1)} x_{12}^{(2)}$$

and

$$P_2 = x_{12}^{(2)} X_1 - x_{12}^{(1)} X_2 + x_{12}^{(1)} x_{11}^{(2)} - x_{11}^{(1)} x_{12}^{(2)}$$

lies in \mathfrak{I}_2 , but $P_1 P_2$ does.

In the proofs of the next two lemmas we modify standard arguments from the PI theory.

Lemma 3.2. *The algebra $\mathfrak{Q}_n/\mathfrak{I}_n$ is isomorphic to the subalgebra of $M_n(\mathcal{C})$ generated by all generic matrices $(x_{ij}^{(k)})$, $k = 1, 2, \dots$, and all λI , $\lambda \in \mathcal{C}$.*

Proof. Let $\Phi : \mathfrak{Q}_n \rightarrow M_n(\mathcal{C})$ be a homomorphism determined by $\Phi(X_k) = (x_{ij}^{(k)})$ and $\Phi(\lambda) = \lambda I$ for $\lambda \in \mathcal{C}$. It is immediate that $\ker \Phi \subseteq \mathfrak{I}_n$. Take $P = P(X_1, \dots, X_m) \in \mathfrak{I}_n$. Thus, $P(A_1, \dots, A_m) = 0$ for all $A_i \in M_n$. Since $\text{char}(F) = 0$, and hence F is infinite, a standard argument shows that $\Phi(P) = 0$. Thus, $\ker \Phi = \mathfrak{I}_n$, and the result follows. \square

The center of $\mathfrak{Q}_n/\mathfrak{I}_n$ is isomorphic to \mathcal{C} , which is a domain. We may therefore form the algebra of central quotients of $\mathfrak{Q}_n/\mathfrak{I}_n$, which consists of elements of the form αR where $R \in \mathfrak{Q}_n/\mathfrak{I}_n$ and α lies in

$$\mathcal{K} := F(x_{ij}^{(k)} \mid 1 \leq i, j \leq n, k = 1, 2, \dots),$$

the field of rational functions in $x_{ij}^{(k)}$ (cf. [12, p. 54]). In order to describe this \mathcal{K} -algebra, we invoke the Capelli polynomials

$$C_{2k-1}(X_1, \dots, X_k, Y_1, \dots, Y_{k-1}) = \sum_{\sigma \in S_k} X_{\sigma(1)} Y_1 X_{\sigma(2)} Y_2 \cdots X_{\sigma(k-1)} Y_{k-1} X_{\sigma(k)}.$$

As it is well-known, C_{2n^2-1} is a polynomial identity of every proper subalgebra of $M_n(E)$ but not of $M_n(E)$ itself, for every field E [12, Theorem 1.4.8].

Lemma 3.3. *The algebra of central quotients of $\mathfrak{Q}_n/\mathfrak{I}_n$ is isomorphic to $M_n(\mathcal{K})$.*

Proof. Since C_{2n^2-1} is not a polynomial identity of $M_n(F)$, it is also not a polynomial identity of the \mathcal{K} -subalgebra of $M_n(\mathcal{K})$ generated by all generic matrices $(x_{ij}^{(k)})$, $k = 1, 2, \dots$. But then this subalgebra is the whole algebra $M_n(\mathcal{K})$. Now we can apply Lemma 3.2. \square

We conclude this section with a small application of Lemma 3.3. Define the image of $P = P(X_1, \dots, X_m) \in \mathfrak{Q}_n$ as

$$\text{im}(P) = \{P(A_1, \dots, A_m) \mid A_1, \dots, A_m \in M_n\}.$$

It is an open question which subsets of M_n can be images of noncommutative polynomials; cf. [8, 14]. It is known and quite easy to see that among linear subspaces of M_n there are only four possibilities: $\{0\}$, the space of all scalar matrices, the space of all trace zero matrices, and M_n . The situation with quasi-polynomials is strikingly different.

Theorem 3.4. *For every linear subspace V of M_n there exists $P \in \mathfrak{Q}_n$ such that $\text{im}(P) = V$.*

Proof. By taking the sums of quasi-polynomials in distinct indeterminates we see that it is enough to prove the theorem for the case where V is one-dimensional, $V = FA$ for some $A \in M_n$. According to Lemma 3.3, we may identify $x_{11}^{(1)}A \in M_n(\mathcal{K})$ with $\lambda^{-1}P_0$ where $0 \neq \lambda \in \mathcal{C}$ and $P_0 \in \mathfrak{Q}_n$. Hence $\text{im}(P_0) \subseteq FA$. Picking an indeterminate $x_{ij}^{(k)}$ of which λ is independent we thus see that $P = x_{ij}^{(k)}P_0$ satisfies $\text{im}(P) = FA$. \square

4. LOCALLY LINEARLY DEPENDENT POLYNOMIALS

Let \mathcal{R} be an F -algebra. Noncommutative polynomials $f_1, \dots, f_t \in F\langle X_1, \dots, X_m \rangle$ are said to be \mathcal{R} -locally linearly dependent if the elements $f_1(r_1, \dots, r_m), \dots, f_t(r_1, \dots, r_m)$ are linearly dependent in \mathcal{R} for all $r_1, \dots, r_m \in \mathcal{R}$. This concept has actually appeared in Operator Theory (see, e.g., [5]), and was recently studied from the algebraic point of view in [4]. The following well-known result (see [2, Theorem 2.3.7] or [12, Theorem 7.6.16]) was used in [4] as an important tool.

Theorem 4.1. *Let \mathcal{R} be a prime algebra. Then $a_1, \dots, a_t \in \mathcal{R}$ are linearly dependent over the extended centroid of \mathcal{R} if and only if $C_{2t-1}(a_1, \dots, a_t, r_1, \dots, r_{t-1}) = 0$ for all $r_1, \dots, r_{t-1} \in \mathcal{R}$.*

By using a similar approach as in the proof of [4, Theorem 3.1], just by applying Theorem 4.1 to the algebra of generic matrices instead of to the free algebra $F\langle \mathcal{X} \rangle$, we get the following characterization of M_n -local linear dependence.

Theorem 4.2. *Noncommutative polynomials f_1, \dots, f_t are M_n -locally linearly dependent if and only if there exist central polynomials c_1, \dots, c_t , not all zero, such that $\sum_{i=1}^t c_i f_i$ is a polynomial identity of M_n .*

Proof. By Theorem 4.1, the condition that f_1, \dots, f_t are M_n -locally linearly dependent is equivalent to the condition that

$$H := C_{2t-1}(f_1, \dots, f_t, Y_1, \dots, Y_{t-1})$$

is a polynomial identity of M_n . Since M_n and the algebra GM_n of $n \times n$ generic matrices satisfy the same polynomial identities, this is the same as saying that H is a polynomial identity of GM_n . Using Theorem 4.1 once again we see that this is further equivalent to the condition that f_1, \dots, f_t , viewed as elements of GM_n , are linearly dependent over the extended centroid of GM_n . Since GM_n is a prime PI-algebra, its extended centroid is the field of fractions of the center of GM_n ; the latter can be identified with central polynomials, and hence the desired conclusion follows. \square

Corollary 4.3. *If noncommutative polynomials f_0, f_1, \dots, f_t are M_n -locally linearly dependent, while f_1, \dots, f_t are M_n -locally linearly independent, then there exist central polynomials c_0, c_1, \dots, c_t , such that $c_0 \neq 0$ and $\sum_{i=0}^t c_i f_i$ is a polynomial identity of M_n .*

5. QUASI-IDENTITIES MULTIPLIED BY CENTRAL POLYNOMIALS

We first treat a very special quasi-identity.

Lemma 5.1. *If $\lambda_i \in \mathcal{C}$, $1 \leq i \leq n^2$, are such that $\sum_{i=1}^{n^2} \lambda_i X_i \in \mathfrak{I}_n$, then each $\lambda_i = 0$.*

Proof. We may assume that $\lambda_i = \lambda_i(X_1, \dots, X_m)$ for some $m \geq n^2$. The set W of all n^2 -tuples $(A_1, \dots, A_{n^2}) \in M_n^{n^2}$ such that $\lambda_i(A_1, \dots, A_{n^2}, T_{n^2+1}, \dots, T_m) = 0$ for all $T_{n^2+1}, \dots, T_m \in M_n$ and all $1 \leq i \leq n^2$ is closed in the Zariski topology of F^{n^4} . Similarly, the set Z of all n^2 -tuples $(A_1, \dots, A_{n^2}) \in M_n^{n^2}$ such that A_1, \dots, A_{n^2} are linearly dependent is also closed - namely, the linear dependence can be expressed through zeros of a polynomial by Theorem 4.1. Of course, $Z \neq M_n^{n^2}$. Suppose that $W \neq M_n^{n^2}$. Then, since F^{n^4} is irreducible (as $\text{char}(F) = 0$), the complements of W and Z in $M_n^{n^2}$ have a nonempty intersection. This means that there exist $A_1, \dots, A_{n^2} \in M_n$ such that $\lambda_i(A_1, \dots, A_{n^2}, T_{n^2+1}, \dots, T_m) \neq 0$ for some $T_{n^2+1}, \dots, T_m \in M_n$, and A_1, \dots, A_{n^2} are linearly independent. However, this is impossible since $\sum_{i=1}^{n^2} \lambda_i X_i \in \mathfrak{I}_n$. Thus, $W = M_n^{n^2}$, i.e., each $\lambda_i = 0$. \square

The condition that $c \in F\langle \mathcal{X} \rangle$ is a central polynomial can be expressed as that there exists $\alpha_c \in \mathcal{C}$ such that $c - \alpha_c \in \mathfrak{I}_n$. Actually, $c - \alpha_c$ is a trace identity since $\alpha_c = \frac{1}{n} \text{tr}(c)$. Therefore the Helling-Procési-Razmyslov theorem yields

Lemma 5.2. *For every central polynomial c of M_n there exists $\alpha_c \in \mathcal{C}$ such that $c - \alpha_c$ is a quasi-identity of M_n contained in the T -ideal of \mathfrak{Q}_n generated by Q_n .*

Theorem 5.3. *For every quasi-identity $P \in \mathfrak{I}_n$ there exists a central polynomial $c \neq 0$ such that cP lies in the T -ideal of \mathfrak{Q}_n generated by Q_n .*

Proof. Let us write

$$P = \sum_{i=1}^{n^2} \lambda_i X_i + \sum \lambda_M M$$

where each M in the second summation is different from X_1, \dots, X_{n^2} . We proceed by induction on the number of summands d in the second summation. If $d = 0$, then $P = 0$ by Lemma 5.1, and the result holds. Let $d > 0$. Pick M_0 such that $M_0 \notin \{X_1, \dots, X_{n^2}\}$ and $\lambda_{M_0} \neq 0$. Note that M_0, X_1, \dots, X_{n^2} are M_n -locally linearly dependent, while X_1, \dots, X_{n^2} are M_n -locally linearly independent. Thus, by Corollary 4.3 there exist central polynomials c_0, c_1, \dots, c_{n^2} such that $c_0 \neq 0$ and

$$f := c_0 M_0 + \sum_{i=1}^{n^2} c_i X_i \in \text{id}(M_n).$$

Lemma 5.2 tells us that for each $i = 0, 1, \dots, n^2$ there exist $\alpha_i \in \mathcal{C}$ such that $c_i - \alpha_i$ is a quasi-identity lying in the T -ideal generated by Q_n . Let us define

$$P' := \alpha_0 P - \lambda_{M_0} \alpha_0 M_0 - \lambda_{M_0} \sum_{i=1}^{n^2} \alpha_i X_i.$$

Writing each α_i as $c_i - (c_i - \alpha_i)$ we see that P' is a quasi-identity. Note that P' involves $d - 1$ summands not lying in $\sum_{i=1}^{n^2} \mathcal{C} X_i$. Therefore the induction assumption yields the existence of a nonzero central polynomial c' such that $c'P'$ lies in the T -ideal generated by Q_n . Setting

$c = c_0 c'$ we thus have $c \neq 0$ and

$$\begin{aligned} cP &= (c_0 - \alpha_0)c'P + \alpha_0 c'P \\ &= (c_0 - \alpha_0)c'P + c'P' + \lambda_{M_0} c' \left(\alpha_0 M_0 + \sum_{i=1}^{n^2} \alpha_i X_i \right) \\ &= (c_0 - \alpha_0)c'P + c'P' - \lambda_{M_0} c' \left((c_0 - \alpha_0)M_0 + \sum_{i=1}^{n^2} (c_i - \alpha_i)X_i \right) + \lambda_{M_0} c' f. \end{aligned}$$

The T-ideal generated by Q_n contains $c_i - \alpha_i$, $0 \leq i \leq n^2$, $c'P'$, as well as f according to Theorem 2.2. Hence it also contains cP . \square

We do not know whether or not Theorem 5.3 gives an optimal conclusion. It does not seem impossible that for some quasi-identities the involvement of a nontrivial central polynomial c is necessary, but we were unable to find an example.

6. MULTILINEAR QUASI-IDENTITIES OF DEGREE n

The multilinearization process works for the quasi-polynomials just as it works for the ordinary noncommutative polynomials. The multilinear quasi-polynomials therefore deserve a special attention. By saying that $P = P(X_1, \dots, X_n)$ is *multilinear* of degree n we mean, of course, that P consists of summands of the form $\lambda(X_{i_1}, \dots, X_{i_k})X_{i_{k+1}} \cdots X_{i_n}$ where $\{1, \dots, n\}$ is the disjoint union of $\{i_1, \dots, i_k\}$ and $\{i_{k+1}, \dots, i_n\}$, and $\lambda(X_{i_1}, \dots, X_{i_k}) \in \mathcal{C}$ is multilinear, i.e., it is a linear combination of monomials of the form $x_{s_1 t_1}^{(i_1)} x_{s_2 t_2}^{(i_2)} \cdots x_{s_k t_k}^{(i_k)}$. A basic example is the Cayley-Hamilton quasi-polynomial Q_n .

Theorem 6.1. *Every multilinear quasi-identity of M_n of degree n is a scalar multiple of Q_n .*

Proof. Let $S_{n,k}$, $1 \leq k \leq n$, denote the set of all permutations $\sigma \in S_n$ such that $\sigma(1) < \sigma(2) < \cdots < \sigma(k)$. For convenience we also set $S_{n,0} = S_n$. Note that a multilinear quasi-polynomial P of degree n can be written as

$$P(X_1, \dots, X_n) = \sum_{k=0}^n \sum_{\sigma \in S_{n,k}} \lambda_{k\sigma}(X_{\sigma(1)}, \dots, X_{\sigma(k)}) X_{\sigma(k+1)} \cdots X_{\sigma(n)}$$

(here, $\lambda_{0\sigma}$ are scalars). By e_{ij} we denote matrix units in M_n .

We assume that P is a quasi-identity, and proceed by a series of claims.

Claim 1. For all $\sigma \in S_{n,k}$, $1 \leq k \leq n$, and all distinct $1 \leq i_1, \dots, i_k \leq n$, we have

$$\lambda_{k\sigma}(e_{11}, \dots, e_{kk}) = \lambda_{k\sigma}(e_{i_1 i_1}, \dots, e_{i_k i_k}) = -\lambda_{k-1, \sigma}(e_{11}, \dots, e_{k-1, k-1}).$$

The proof is by induction on k . First, take $0 \leq j \leq n-1$ and substitute

$$e_{1+j, 1+j}, e_{1+j, 2+j}, e_{2+j, 3+j}, \dots, e_{n-1+j, n+j}$$

(with addition modulo n) for $X_{\sigma(1)}, \dots, X_{\sigma(n)}$ in P . Considering the coefficient at $e_{1+j, n+j}$ we get $\lambda_{0\sigma} + \lambda_{1\sigma}(e_{1+j, 1+j}) = 0$. This implies the truth of Claim 1 for $k = 1$. Let $k > 1$ and take $\sigma \in S_{n,k}$. Choose a subset of $\{1, \dots, n\}$ with $k-1$ elements, $\{i_{n-k+2}, \dots, i_n\}$, and let $\{j_1, \dots, j_{n-k+1}\}$ be its complement. Let us substitute

$$e_{i_{n-k+2}, i_{n-k+2}}, \dots, e_{i_n, i_n}, e_{j_1, j_1}, e_{j_1, j_2}, e_{j_2, j_3}, \dots, e_{j_{n-k}, j_{n-k+1}}$$

for $X_{\sigma(1)}, \dots, X_{\sigma(n)}$, respectively, in P . Similarly as above, this time by considering the coefficient at $e_{j_1, j_{n-k+1}}$, we obtain

$$\lambda_{k-1, \sigma}(e_{i_{n-k+2}, i_{n-k+2}}, \dots, e_{i_n, i_n}) + \lambda_{k\sigma}(e_{i_{n-k+2}, i_{n-k+2}}, \dots, e_{i_n, i_n}, e_{j_1, j_1}) = 0.$$

The desired conclusion follows from the induction hypothesis.

Claim 2. For all $\sigma, \tau \in S_{n,k}$, $0 \leq k \leq n-1$, and all distinct $1 \leq i_1, \dots, i_k \leq n$, we have

$$\lambda_{k\sigma}(e_{11}, \dots, e_{kk}) = \lambda_{k\tau}(e_{i_1 i_1}, \dots, e_{i_k i_k}).$$

Evaluating P at e_{11}, \dots, e_{nn} results in

$$\lambda_{n-1, \sigma_i}(e_{11}, \dots, e_{i-1, i-1}, e_{i+1, i+1}, \dots, e_{nn}) = \lambda_{n-1, \text{id}}(e_{11}, \dots, e_{n-1, n-1})$$

for all $1 \leq i \leq n-1$, where σ_i stands for the cycle $(i \ i+1 \ \dots \ n)$. Accordingly, since $S_{n, n-1}$ consists of id and all σ_i , $1 \leq i \leq n-1$, the case $k = n-1$ follows by Claim 1. We may now assume that $k < n-1$ and that Claim 2 holds for $k+1$. Take $\sigma \in S_{n,k}$. If $\sigma \in S_{n, k+1}$ then

$$\lambda_{k\sigma}(e_{11}, \dots, e_{kk}) = -\lambda_{k+1, \sigma}(e_{11}, \dots, e_{k+1, k+1})$$

by Claim 1. If $\sigma \notin S_{n, k+1}$ there exists $1 \leq i \leq k$ such that $\sigma(k+1) < \sigma(i)$. Substituting

$$e_{11}, \dots, e_{kk}, e_{k+1, k+1}, e_{k+1, k+2}, e_{k+2, k+3}, \dots, e_{n-1, n}$$

for $X_{\sigma(1)}, \dots, X_{\sigma(n)}$ in P we infer that there for a certain permutation τ (specifically, $\tau = \sigma \circ (k+1 \ k \ \dots \ i+1 \ i)$) we have

$$\lambda_{k\sigma}(e_{11}, \dots, e_{kk}) = -\lambda_{k+1, \tau}(e_{11}, \dots, e_{i-1, i-1}, e_{k+1, k+1}, e_{i+1, i+1}, \dots, e_{k-1, k-1}).$$

Since every $\lambda_{k\sigma}(e_{11}, \dots, e_{kk})$ is associated to an evaluation of $\lambda_{k+1, \tau}$, Claim 2 follows by the induction hypothesis and Claim 1.

Claim 3. $P = \lambda_{0, \text{id}} Q_n$.

By Claim 2 we have $\lambda_{0\sigma} = \lambda_{0\tau}$ for all $\sigma, \tau \in S_n$. Accordingly, $R := P - \lambda_{0, \text{id}} Q_n$ does not involve summands of the form $\mu X_{\sigma(1)} \dots X_{\sigma(n)}$, $\mu \in F$, and can be therefore written as

$$R(X_1, \dots, X_n) = \sum_{k=1}^n \sum_{\sigma \in S_{n,k}} \mu_{k\sigma}(X_{\sigma(1)}, \dots, X_{\sigma(k)}) X_{\sigma(k+1)} \dots X_{\sigma(n)}$$

We must prove that $R = 0$, i.e., each $\mu_{k\sigma} = 0$. We proceed by induction on k . For $k = 0$ this holds by the hypothesis, so let $k > 0$. It suffices to show that $\mu_{k\sigma}(e_{i_1 j_1}, \dots, e_{i_k j_k}) = 0$ for arbitrary matrix units $e_{i_1 j_1}, \dots, e_{i_k j_k}$. Choose distinct l_1, \dots, l_{n-k} such that $l_s \neq i_t$ for all s, t . Substitute

$$e_{i_1 j_1}, \dots, e_{i_k j_k}, e_{l_1 l_2}, e_{l_2 l_3}, \dots, e_{l_{n-k} i_1}$$

for $X_{\sigma(1)}, \dots, X_{\sigma(n)}$ in P . There is only one way to factorize $e_{l_1 i_1}$ as a product of at most $n-k$ chosen matrix units, i.e., $e_{l_1 i_1} = e_{l_1 l_2} e_{l_2 l_3} \dots e_{l_{n-k} i_1}$. By induction hypothesis it thus follows that $\mu_{k\sigma}(e_{i_1 j_1}, \dots, e_{i_k j_k}) = 0$. \square

7. QUASI-IDENTITIES IN ONE INDETERMINATE

The set of quasi-polynomials in one indeterminate, i.e., those of the form

$$p(X) = \sum_{i=0}^m \lambda_i(X) X^i,$$

forms a commutative subalgebra of \mathfrak{Q}_n . Describing quasi-identities in one indeterminate is a rather easy task. We first record a lemma without a proof since it is practically the same as that of Lemma 5.1. The crucial part is observing that the set of all $A \in M_n$ such that I, A, \dots, A^{n-1} are linearly dependent is, as follows from Theorem 4.1, closed with respect to the Zariski topology. On the other hand, this lemma also follows from [13, Lemma 14.7].

Lemma 7.1. *If $\lambda_i \in \mathcal{C}$, $1 \leq i \leq n-1$, are such that $\sum_{i=1}^{n-1} \lambda_i X^i \in \mathfrak{I}_n$, then each $\lambda_i = 0$.*

Theorem 7.2. *If a quasi-polynomial in one indeterminate $p(X)$ is a quasi-identity of M_n , then there exists a quasi-polynomial $r(X)$ such that $p(X) = r(X)q_n(X)$.*

Proof. Let $p(X) = \sum_{i=0}^m \lambda_i(X) X^i$. The proof is by induction on m . In view of Lemma 7.1 we may assume that $m \geq n$. Note that $p(X) - \lambda_m(X) X^{m-n} q_n(X)$ is a quasi-identity for which the induction assumption is applicable. Therefore $p(X) - \lambda_m(X) X^{m-n} q_n(X) = r_1(X) q_n(X)$ for some $r_1(X)$, and hence $p(X) = (\lambda_m(X) X^{m-n} + r_1(X)) q_n(X)$. \square

This theorem implies that $q_n(h(X))$, where $h(X)$ is any quasi-polynomial in one indeterminate, can be written as $r(X)q_n(X)$ for some $r(X)$. Of course, there are other ways to establish this; at any rate, however, this proof is very short.

8. OTHER FUNCTIONAL IDENTITIES

Let R be a ring with center C and let $E_1, \dots, E_{m+1} : R^m \rightarrow R$ be functions satisfying

$$\sum_{i=1}^{m+1} E_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{m+1}) x_i \in C \quad \text{for all } x_1, \dots, x_{m+1} \in R.$$

This is one of the most important examples of functional identities. Its standard solution is defined simply as $E_i = 0$ for every i . It turns out that these are also the only possible solutions in a rather large class of rings. Let us, however, confine ourselves to matrices: If $n > m+1$ then this functional identity has only standard solutions on M_n [3, Lemma 2.8, Corollary 2.23], while for $n \leq m+1$ the Cayley-Hamilton quasi-polynomial Q_n readily gives rise to nonstandard solutions. Since Q_n is symmetric, the basic examples of these nonstandard solutions are such that all the E_i 's are equal, $E_i = E$ for $1 \leq i \leq m+1$, and multilinear. Let us, therefore, consider such a functional identity. By taking all the x_i 's to be equal, we can write it as $E(x, \dots, x)x \in C$.

Recall that by τ_i we have denoted the coefficient functions of q_n . Define a quasi-polynomial $p_n(X)$ by

$$p_n(X) = X^{n-1} + \tau_1(X) X^{n-2} + \dots + \tau_{n-1}(X),$$

so that $q_n(X) = X p_n(X) + \tau_n(X)$.

Theorem 8.1. *If $E : M_n^m \rightarrow M_n$, $n \leq m+1$, is a multilinear map such that $E(A, \dots, A)A \in F$ for all $A \in M_n$, then there exists a map $\mu : M_n \rightarrow F$ such that $E(A, \dots, A) = \mu(A)p_n(A)$ for all $A \in M_n$.*

Proof. Note, first of all, that without loss of generality we may assume that E is symmetric.

We claim that for each $0 \leq j \leq m$ there is a multilinear quasi-polynomial $H_j(X_1, \dots, X_j)$ such that $H_j(A_1, \dots, A_j) = E(A_1, \dots, A_j, I, \dots, I)$ for all $A_1, \dots, A_j \in M_n$. Since $E(I, \dots, I)$ obviously lies in F , this is trivially true for $j = 0$. We may therefore assume that it is true for all positive integers smaller than j . Taking $A_{j+1} = \dots = A_{m+1} = I$ in the multilinear version of our functional identity,

$$\sum_{i=1}^{m+1} E(A_1, \dots, A_{i-1}, A_{i+1}, \dots, A_{m+1})A_i \in F \quad \text{for all } A_1, \dots, A_{m+1} \in M_n,$$

we obtain

$$E(A_2, \dots, A_j, I, \dots, I)A_1 + \dots + E(A_1, \dots, A_{j-1}, I, \dots, I)A_j + (m+1-j)E(A_1, \dots, A_j, I, \dots, I) \in F$$

Since $E(A_2, \dots, A_j, I, \dots, I), \dots, E(A_1, \dots, A_{j-1}, I, \dots, I)$ can be expressed through multilinear quasi-polynomials by assumption, it follows that the same holds for $E(A_1, \dots, A_j, I, \dots, I)$. Our claim is thus proved.

For $j = m$ we have $H_m(A_1, \dots, A_m) = E(A_1, \dots, A_m)$, and hence

$$E(A, \dots, A) = H_m(A, \dots, A) = \lambda_0(A)A^m + \lambda_1(A)A^{m-1} + \dots + \lambda_{m-1}(A)A + \lambda_m(A),$$

where the functions $\lambda_i : M_n \rightarrow F$ arise from multilinear maps and can be therefore expressed as polynomial functions in the entries of A (and λ_0 is a scalar). By the Cayley-Hamilton theorem, each A^j , $j \geq n$, can be written as $\sum_{i=1}^{n-1} \alpha_i(A)A^i$ where the α_i 's are also polynomial functions in the entries of A . Accordingly, we have

$$E(A, \dots, A) = \mu_0(A)A^{n-1} + \mu_1(A)A^{n-2} + \dots + \mu_{n-2}(A)A + \mu_{n-1}(A),$$

where the μ_i 's can be considered as elements in \mathcal{C} . Now, using $E(A, \dots, A)A \in F$ together with

$$A^n = -(\tau_1(A)A^{n-1} + \dots + \tau_{n-1}(A)A + \tau_n(A)),$$

it follows that

$(\mu_1(A) - \mu_0(A)\tau_1(A))A^{n-1} + (\mu_2(A) - \mu_0(A)\tau_2(A))A^{n-2} + \dots + (\mu_{n-1}(A) - \mu_0(A)\tau_{n-1}(A))A \in F$ for every $A \in M_n$. Using Lemma 7.1 we may now conclude that $\mu_i(A) = \mu_0(A)\tau_i(A)$, $i = 1, \dots, n-1$. Consequently, $E(A) = \mu(A)p_n(A)$ where $\mu = \mu_0$. \square

We conclude the paper with a simple example which indicates that finding all nonstandard solutions of functional identities of the type $\sum_i E_i x_i \in C$ on M_n may be a very difficult problem.

Example 8.2. Define $E : M_2 \rightarrow M_2$ by $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mapsto \begin{bmatrix} a_{22} & a_{12} \\ -a_{21} & -a_{11} \end{bmatrix}$. Then $E(A)B - E(B)A \in F$ for all $A, B \in M_2$.

Acknowledgement. The authors would like to thank Igor Klep for insightful comments.

REFERENCES

- [1] K. I. Beidar, M. A. Chebotar, On functional identities and d -free subsets of rings II, *Comm. Algebra* **28** (2000), 3953–3972.
- [2] K. I. Beidar, W. S. Martindale 3rd, A. V. Mikhalev, *Rings with generalized identities*, Marcel Dekker, Inc., 1996.
- [3] M. Brešar, M. A. Chebotar, W. S. Martindale 3rd, *Functional identities*, Birkhäuser Verlag, 2007.
- [4] M. Brešar, I. Klep, A local-global principle for linear dependence of noncommutative polynomials, *Israel J. Math.*, to appear.
- [5] J. F. Camino, J. W. Helton, R. E. Skelton, J. Ye, Matrix inequalities: a symbolic procedure to determine convexity automatically, *Int. Eq. Oper. Th.* **46** (2003) 399–454.
- [6] M. A. Chebotar, W.-F. Ke, P.-H. Lee, R. Zhang, On maps preserving zero Jordan products, *Monatsh. Math.* **149** (2006), 91–101.
- [7] H. Helling, Eine Kennzeichnung von Charakteren auf Gruppen und assoziativen Algebren, *Comm. Algebra* **1** (1974), 491–501.
- [8] A. Kanel-Belov, S. Malev, L. H. Rowen, The images of non-commutative polynomials evaluated on 2×2 matrices, *Proc. Amer. Math. Soc.* **140** (2012), 465–478.
- [9] A. Kanel-Belov, L. H. Rowen, *Computational aspects of polynomial identities*, A K Peters, 2005.
- [10] C. Procesi, The invariant theory of $n \times n$ matrices, *Adv. Math.* **19** (1976), 306–381.
- [11] Yu. P. Razmyslov, Identities with trace in full matrix algebras over a field of characteristic zero, *Izv. Akad. Nauk SSSR Ser. Mat.* **38** (1974), 723–756.
- [12] L. H. Rowen, *Polynomial identities in ring theory*, Academic Press, 1980.
- [13] D. J. Saltman, *Lectures on division algebras*, Amer. Math. Soc., 1999.
- [14] Š. Špenko, On the image of a noncommutative polynomial, *J. Algebra*, to appear.

M. BREŠAR, FACULTY OF MATHEMATICS AND PHYSICS, UNIVERSITY OF LJUBLJANA, AND FACULTY OF NATURAL SCIENCES AND MATHEMATICS, UNIVERSITY OF MARIBOR, SLOVENIA
E-mail address: `matej.bresar@fmf.uni-lj.si`

Š. ŠPENKO, INSTITUTE OF MATHEMATICS, PHYSICS, AND MECHANICS, LJUBLJANA, SLOVENIA
E-mail address: `spela.spenko@imfm.si`